<u>Homework Assignment-7</u> 2 rt TU 88

<u>POM 500 Statistical Analysis</u>

Note: Attempt all questions as per rubric. Problems including case study has a weightage of 10 marks each. The maximum you can score is 50. <u>Use Excel function wherever possible</u>.

<u>Problem-1</u>

According to the Hospital Care cost Institute, the annual expenditure for prescription drugs is $838 per person in the Northeast region of the country. A sample of 60 individuals in the Midwest shows a per person annual expenditure for prescription drugs of $745. Use a population standard deviation of $300 to answer the following questions.

a) Formulate hypotheses for a test to determine whether the sample data support the conclusion that the population annual expenditure for prescription drugs per person is lower in the Midwest than in the Northeast.

b) What is the value of the test statistic?

c) What is the p-value?

d) At = 0.01, what is your conclusion?

. Formulate hypotheses for a testtestl

- $\bar{X}$ is the sample mean ($745)
- μ is the population mean ($838)
- σ is the population standard deviation ($300)
- n is the sample size (60) low

Substituting the given values into the formula, we get:

z = ($745 - $838) / ($300 / √60) ≈ -3.16

So, the value of the test statistic is approximately -3.16.

3. P-value

The p-value is the probability of observing a test statistic as extreme as the one calculated, assuming the null hypothesis is true. For a z-score of -3.1p6, the p-value can be found using a standard normal distribution table or a statistical software. The exact p-value is less than 0.01.

4. Conclusion at α = 0.01

At a significance level of $\alpha = 0.01$, if the p-value is less than α, we reject the null hypothesis. Since our calculated p-value is less than 0.01, we reject the null hypothesis and conclude that the data provides sufficient evidence to support the claim that the population annual expenditure for prescription drugs per person is lower in the Midwest than in the Northeast.

e) Problem-2

The United States ranks ninth in the world in per capita chocolate consumption; Forbes reports that the average American eats 9.5 pounds of chocolate annually. Suppose you are curious whether chocolate consumption is higher in Hershey, Pennsylvania, the location of the Hershey Company’s corporate headquarters. A sample of 36 individuals from the Hershey area showed a sample mean annual consumption of 10.05 pounds and a standard deviation of s= 1.5 pounds. Using =.05, do the sample results support the conclusion that mean annual consumption of chocolate is higher in Hershey than it is throughout the United States?

To answer your question, we need to conduct a one-sample t-test. This test is used to determine whether a sample mean significantly differs from a known or hypothesized population mean.

Here are the steps:

Step 1: State the Hypotheses

The first step in hypothesis testing is to set a null hypothesis and an alternative hypothesis.

- Null hypothesis (H0): $\mu = 9.5$ (Mean chocolate consumption in Hershey is the same as the national average)
- Alternative hypothesis (H1): $\mu > 9.5$ (Mean chocolate consumption in Hershey is greater than the national average)

Step 2: Calculate the Test Statistic

The test statistic for a one-sample t-test is calculated as follows:

$t = (\bar{X} - \mu) / (s / \sqrt{n})$

Where:

- $\bar{X}$ is the sample mean

- μ is the hypothesized population mean
- s is the sample standard deviation
- n is the sample size

Substituting the given values:

$t = (10.05 - 9.5) / (1.5 / \sqrt{36}) = 2.2$

Step 3: Determine the Critical Value

The critical value for a one-tailed t-test with 35 degrees of freedom (n-1) and $\alpha = 0.05$ can be found in a t-distribution table or calculated using statistical software. The critical value is approximately 1.689.

Step 4: Make the Decision

If the absolute value of the test statistic is greater than the critical value, we reject the null hypothesis.

In this case, $2.2 > 1.689$, so we reject the null hypothesis.

Conclusion

The sample results support the conclusion that the mean annual consumption of chocolate is higher in Hershey, Pennsylvania than it is throughout the United States.

Problem-3

Last year, a soft drink manufacturer had 21% of the market. In order to increase their

To answer your question, we need to conduct a one-sample t-test. This test is used to determine whether a sample mean significantly differs from a known or hypothesized population mean.

1. Set up the null and the alternative hypotheses.

1. Determine the test statistic.

1. Determine the p-value.

1. Using α = .05, test to determine if more than 21% of the population will like the new soft drink.

Case Study: Quality Associates, Inc.

Quality Associates, Inc. Quality Associates, Inc., a consulting firm, advises its clients about sampling and statistical procedures that can be used to control their manufacturing processes. In one particular application, a client gave Quality Associates a sample of 800 observations taken during a time in which that client's process was operating satisfactorily. The sample standard deviation for these data was 0.21; hence, with so much data, the population standard deviation was assumed to be 0.21. Quality Associates then suggested that random samples of size 30 be taken periodically to monitor the process on an ongoing basis. By analyzing the new samples, the client could quickly learn whether the process was operating satisfactorily. When the process was not operating satisfactorily, corrective action could be taken to eliminate the problem. The design specification indicated the mean for the process should be 12. The hypothesis test suggested by Quality Associates follows.

H_0: $\mu = 12$

H_a: $\mu \neq 12$

Corrective action will be taken any time H_0 is rejected.

The following samples were collected at hourly intervals during the first day of operation of the new statistical process control procedure. These data are available in the data set *Quality*.

Managerial Report

1. Conduct a hypothesis test for each sample at the 0.01 level of significance and determine what action, if any, should be taken. Provide the test statistic and p-value for each test.

2. Compute the standard deviation for each of the four samples. Does the assumption of 0.21 for the population standard deviation appear reasonable?

3. Compute limits for the sample mean $\bar{x}$ around $\mu=12$ such that, as long as a new sample mean is within those limits, the process will be considered to be operating satisfactorily. If $\bar{x}$ exceeds the upper limit or if $\bar{x}$ is below the lower limit, corrective action will be taken. These limits are referred to as upper and lower control limits for quality control purposes.

4. Discuss the implications of changing the level of significance to a larger value. What mistake or error could increase if the level of significance is increased? Data set *Quality*

Sample 1	Sample 2	Sample 3	Sample 4
11.55	11.62	11.91	12.02
11.62	11.69	11.36	12.02
11.52	11.59	11.75	12.05
11.75	11.82	11.95	12.18
11.90	11.97	12.14	12.11
11.64	11.71	11.72	12.07
11.80	11.87	11.61	12.05
12.03	12.10	11.85	11.64
11.94	12.01	12.16	12.39
11.92	11.99	11.91	11.65
12.13	12.20	12.12	12.11
12.09	12.16	11.61	11.90
11.93	12.00	12.21	12.22
12.21	12.28	11.56	11.88
12.32	12.39	11.95	12.03
11.93	12.00	12.01	12.35
11.85	11.92	12.06	12.09
11.76	11.83	11.76	11.77
12.16	12.23	11.82	12.20
11.77	11.84	12.12	11.79
12.00	12.07	11.60	12.30
12.04	12.11	11.95	12.27
11.98	12.05	11.96	12.29
12.30	12.37	12.22	12.47
12.18	12.25	11.75	12.03
11.97	12.04	11.96	12.17
12.17	12.24	11.95	11.94
11.85	11.92	11.89	11.97
12.30	12.37	11.88	12.23
12.15	12.22	11.93	12.25

The interval for population mean is (11.8764 ,12.0410)

b. Sample 2:

$\acute{x}-t$

$11.9464 < \mu < 12.1110$
The interval for population mean is (11.9464,12.1110)
c. Sample 3:
$\acute{x} - t$
$11.8116 < \mu < 11.9664$
The interval for population mean is (11.8116,11.9664)
d. Sample 4:
$\acute{x} - t$
$12.0043 < \mu < 12.1583$
The interval for population mean is (12.0043,12.1583)
4. Discuss the implications of changing the level of significance to a larger value. What mistake or
error could increase if the level of significance is increased?
We can see that, the margin of error will increase and the width of confidence interval will
decrease with the increase in level of significance.